

基于迁移学习和Xception网络的混凝土裂纹识别研究

徐峥匀¹, 钱松荣^{1,2}

(1. 贵州大学机械工程学院, 贵州 贵阳 550025;

2. 贵州大学公共大数据国家重点实验室, 贵州 贵阳 550025)

✉xu.zhengyun@qq.com; 909718747@qq.com



摘要: 针对深层卷积神经网络检测表面结构裂纹耗时长、精度不够高的问题, 基于Xception网络进行自适应调整重构其分类器, 利用图像扩增技术扩充数据集后, 引入迁移学习的方法对Xception网络进行训练。同时, 与构建的ResNet50、InceptionV3和VGG19三个深层卷积神经网络模型进行对比实验, 重新验证其性能。实验证明, 引入迁移学习不仅可以提升模型的整体性能, 还能缩减训练深层卷积神经网络的时间, 训练的模型在数据集上的识别精确率达到96.24%, 在对比实验中达到96.50%。

关键词: 迁移学习; 卷积神经网络; 图像识别; 图像扩增

中图分类号: TP399 **文献标识码:** A

Research on Concrete Crack Identification based on Transfer Learning and Xception

XU Zhengyun¹, QIAN Songrong^{1,2}

(1. College of Mechanical Engineering, Guizhou University, Guiyang 550025, China;

2. State Key Laboratory of Public Big Data, Guizhou University, Guiyang 550025, China)

✉xu.zhengyun@qq.com; 909718747@qq.com

Abstract: Aiming at the problems of time consuming and low accuracy of deep convolution neural network in detecting surface structural cracks, this paper proposes to adaptively adjust and reconstruct its classifier based on Xception, and use image augmentation technology to expand the dataset. After that, transfer learning method is introduced for training Xception. At the same time, its performance is re-verified by comparing it with three deep convolution neural network models, ResNet50, InceptionV3 and VGG19. Experiments show that the introduction of transfer learning not only improves the overall performance of the model, but also reduces the time of training the deep convolutional neural network. Recognition accuracy of the trained model on the dataset reaches 96.24%, and 96.50% in the comparative experiment.

Keywords: transfer learning; convolutional neural network; image recognition; image augmentation

1 引言(Introduction)

桥梁结构会随着时间的推移而退化, 从而出现裂纹甚至坍塌, 造成安全事故^[1]。深度学习^[2](Deep Learning)技术处理大量数据和自动化特征提取的能力使其能够开发混凝土结构裂缝检测问题的解决方案。而迁移学习^[3](Transfer Learning)可以从相关或者相近领域中转移已经标注好的数据、参数或者知识结构, 从而完成或改进目标领域或目标任务的学习效果。

本文基于迁移学习和Xception^[4]网络, 使用混凝土裂纹图像数据集训练出能够检测裂纹的模型, 利用图像扩增技术扩充数据集; 比较了使用扩增数据集和迁移学习前后模型在测试集上的表现, 得到一个训练时间短并且具有较高精度的网

络模型。

2 实验流程(Experimental process)

基于迁移学习的Xception网络裂纹识别模型建模流程如图1所示。首先, 对原有数据集的图片进行图片增强, 从而增加训练集; 然后将数据分类并对不同种类进行标记, 并且对其进行标准化或归一化处理, 以消除数据特征因数量级不同而对训练过程产生的影响; 接着, 引入迁移学习的方法将相关领域中的知识结构和训练参数迁移到构建好的模型当中进行训练, 训练过程中对模型进行微调以提高其精确度; 最后, 使用测试数据集在所得模型上进行验证, 检测结果以指定的评价指标来表示。实验中使用到的方法和构建的模型将

在下文介绍。

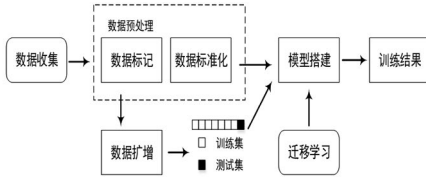


图1 方法流程图

Fig.1 Method flow chart

3 实验搭建(Experimental construction)

3.1 数据来源及处理

目前还没有发布统一的桥梁裂缝数据库。因此,为了满足实验要求,基于网上的共享数据集^[5](<https://github.com/tjdxxy/crack-detection>)进行实验。原始数据为Phantom 4 Pro的CMOS表面阵列相机采集的2,068张裂纹图像,分辨率为 1024×1024 。实验中所使用的桥梁裂纹数据集是基于原始数据集处理后得到的,分别为裂纹图像4,058张和背景图像2,011张,占比约为2:1。数据集中的部分图片如图2所示,其中上方一组是含有裂纹的图像,下方一组为无裂纹(正常)图像。

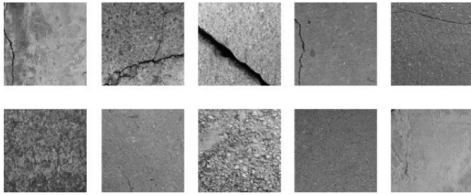


图2 原始数据集图像示例

Fig.2 Raw dataset image example

数据的标准化和归一化^[6]可以消除因特征不同、数量级相差较大带来的影响,同时可以使数据保持原始分布,对提高模型的精度具有积极影响。将图片规格尺寸调整为 $256 \times 256 \times 3$ 。使用图像增广技术^[7]对训练图像做一系列随机改变,以产生相似但不相同的训练样本,这样随机改变训练样本可以降低模型对某些属性的依赖性。

图像增广的具体操作如下:

(1)裁剪和翻转:通过对图像进行不同方式的裁剪和翻转,使得感兴趣的物体或者特征出现在不同位置,这种方法可以减轻模型对数据集中裂纹出现位置的依赖性。

(2)变化颜色:通过调节整体亮度、色彩等因素,起到模拟不同天气环境的作用,以此来降低模型对光照色彩的敏感度。

实际操作时,将几种效果进行随机组合,利用图像增广技术的一个示例如图3所示,上方图像为原始数据集中任意一张图像,下方图像是其经过四次图像增广技术分别得到的图像。最终数据集被扩增至25,576张,其中裂纹图像16,642张,无裂纹图像8,934张。对所有图像进行标记,根据有无裂纹分别标记为正负样本。

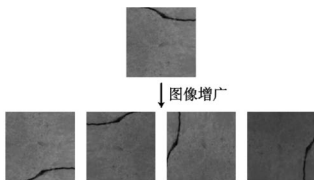


图3 图像增广示例

Fig.3 Image augmentation example

归一化处理如下:

在图像输入网络之前,对其进行归一化处理,将图像的特征数据通过等式(1)进行变换:

$$x_i^* = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (1)$$

式(1)中的 $x = (x_1, x_2, \dots, x_m)$,表示图像中每个像素点的值所组成的集合; $x_i \in x$; $\max(x)$ 和 $\min(x)$ 分别表示为该集合中像素值的最大取值和最小取值; x_i^* 表示通过计算后将任意取值范围内的 x_i 值转化为 $[0, 1]$ 区间内的值,即为归一化处理后的特征值。

3.2 实验平台

实验所选择的编写语言为Python^[8]。Python作为开源软件,不仅有强大的数据处理和算法模型搭建能力,也因拥有丰富而强大的库被广泛应用于科研中。其中Numpy库封装了许多基础的矩阵和向量的操作,而Scipy库则在此基础上提供了更丰富的功能,比如各种统计常用的分布和算法都能迅速地在Scipy库中找到,加上Matplotlib库用来提供数据可视化,这些方法被高效应用于实验中。

TensorFlow^[9]是一个采用数据流图(Data Flow Graphs),用于数值计算的开源软件库,而现在已经发展成为一个平台,用于机器学习和深度学习方面的研究,在易用性、分布式训练和部署等方面都取得了长足的进步。实际上,TensorFlow已进化成为一个更全面的“工具、库和社区资源生态系统”^[9],研究人员可以很方便地使用它来进行各种实验。TensorFlow 2.0生态系统包含对Python的支持,本文中的实验均使用其框架搭建模型进行训练。

3.3 模型建立

(1)Xception网络模型

深层的卷积神经网络对复杂图像的识别相比于浅层结构往往能够取得更好的效果,本文中使用的Xception网络是谷歌公司继Inception系列网络后的改进结构,其主要原理是采用深度可分离卷积来替换原来网络中的卷积操作,在基本不增加网络复杂度的同时,进一步提高了模型的效果。

为调整成适应本次实验的结构,将网络的输入调整为 $256 \times 256 \times 3$,并对该网络只选取卷积层结构,对后面的输出层进行自定义搭建,搭建完成以后的模型结构示意图如图4所示。在图4中可以看到,输入层图片调整尺寸后,批量输入Xception网络的卷积层结构中,经过全局池化,连续经过三个自定义的全连接层^[10]后得到输出结果,其中全连接层的参数(神经元)个数分别为512、256和64。前两个全连接层分别连接参数大小相同的Dropout层^[11],训练时每传递一次数据,Dropout层就会随机选择并删除神经元信号,在模型中删除的比例均设置为20%,用于抑制模型的过拟合^[12]。

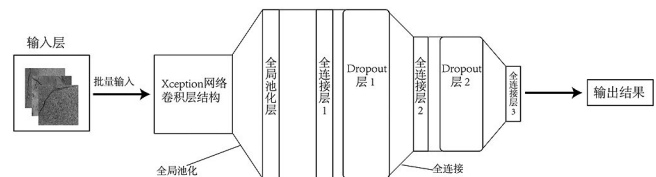


图4 模型结构示意图

Fig.4 Schematic diagram of model structure

(2) 基于迁移学习理论的模型构建

在给定源域 N_s 和学习任务 W_s 、目标域 N_t 和学习任务 W_t 时，迁移学习的目的是获取源域 N_s 和学习任务 W_s 中的知识以帮助提高预测函数 $f_t(\bullet)$ 的学习，其中 $N_s \neq N_t$ 或者 $W_s \neq W_t$ 。其基本原理如图5所示。

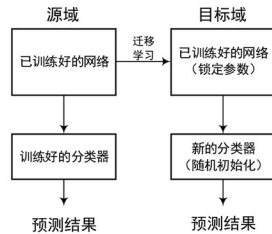


图5 迁移学习示意图

Fig.5 Schematic diagram of transfer learning

预训练网络是一个已经在其他数据集(主要是大规模的图像识别数据集)上训练好的卷积神经网络，所以训练的数据集往往非常大并且具有通用性，网络学习提取到的特征空间层次结构能够用作特征提取模型。学习到的特征在不同问题之间往往是可转移的，使得深度学习和迁移学习相比于其他方法更具有优势，它使得深度学习对于小数据问题非常有效。实验中使用到的源域数据集ImageNet^[13]是一个已经标记好种类的图像数据库(常用于机器视觉研究)，目前已经超过20,000个类别。在图5中，保持原本预先训练好的网络中的卷积基参数不变，将其冻结，删除其训练好的分类器(图5后半部分)并且创建新的分类器，对其参数随机初始化，从而搭建出新的模型并使用数据集进行训练，得到新的预测模型。

3.4 评价指标

本文主要采用损失值(Loss)、均方根误差(RMSE)和精确率(Precision)作为模型性能的评价指标。

RMSE值的计算方法如下：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2} \quad (2)$$

式(2)中的 y_i 和 y_i^* 分别代表第 i 个样本的真实值和预测值。

损失值是衡量卷积神经网络性能“劣化程度”的一个指标，即当前模型与监测数据的不匹配程度以及不一致程度。通过最大化地减少或降低损失函数，可以减小模型预测误差，使得模型逐渐收敛。本文采用二元交叉熵进行损失值计算，其公式如下：

$$f_p(q) = -\frac{1}{n} \sum_{i=1}^n y_i \cdot \ln p(y_i) + (1 - y_i) \cdot \ln (1 - p(y_i)) \quad (3)$$

其中， y_i 代表第 i 个样本的预测值，正值返回1，负值返回0； $p(y)$ 是所有样本的预测正概率。

精确率的计算公式如下：

$$p = \frac{TP}{TP+FP} \quad (4)$$

在式(4)中， TP 代表模型预测值为正例，真实值也是正例的数目； FP 表示模型预测值为正例，而真实值是负例的数目。

损失值和RMSE值下降得越低、精确率越高，那么模型的性能越好；反之，模型性能越差。

4 实验结果及分析(Experimental results and analysis)

实验采用小批量随机梯度下降法，将扩增后不同类型的数据分别打乱，其中分别选出80%的数据作为训练数据，其余部分作为测试数据。数据集划分表如表1所示。为了得到稳定有效的结果，分别进行四类模型的实验：(1)仅使用原始数据集；(2)使用扩增数据集；(3)使用迁移学习；(4)同时使用扩增数据集和迁移学习。

表1 数据集划分表

Tab.1 Dataset partition table

类型	训练集数	测试集数	总计
无裂纹	7,147	1,787	8,934
有裂纹	13,314	3,328	16,642
总计	20,461	5,115	25,576

实验过程中使用相同的优化器(Adam)^[14]和相同的学习率 1×10^{-4} ，训练过程总共10个轮次，对每个模型进行10次交叉验证后，获得了最稳定的结果。表2显示了损失值、RMSE值和精确率的最终实验结果，图6、图7和图8分别显示了所有模型训练过程中损失值、RMSE值和精确率的变化趋势。

表2 实验结果

Tab.2 Experimental results

模型	损失值	RMSE值	精确率/%
原始数据集	0.0880	0.1712	79.82
扩增数据集	0.0714	0.1325	84.67
迁移学习	0.0340	0.0574	96.24
两种方法结合	0.0326	0.0520	97.13

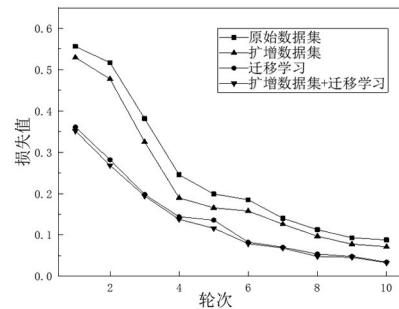


图6 损失值变化趋势图

Fig.6 Variation trend of loss value

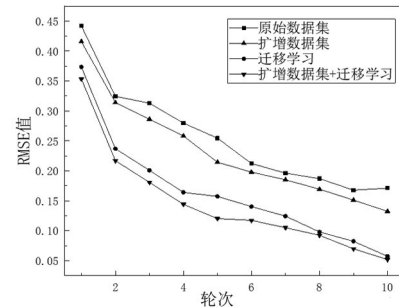


图7 RMSE值变化趋势图

Fig.7 Variation trend of RMSE value

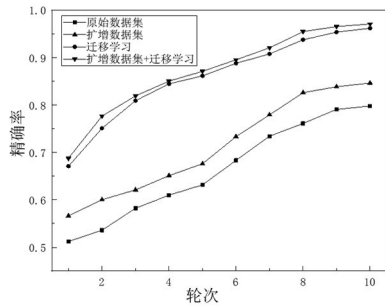


图8 精确率变化趋势图

Fig.8 Variation trend of precision rate

为了进一步检验网络的识别能力,本文利用相同的方式构建了ResNet50、InceptionV3和VGG19三个深度卷积神经网络进行对比实验,同样扩增数据集并引入迁移学习对四个网络重新进行测试,使用精确率、召回率和 F_1 指数作为评价指标,所得结果如表3所示。

表3 对比实验结果

Tab.3 Results of comparative experiments

模型	精确率/%	召回率/%	F_1 指数/%
Xception	96.50	96.78	96.64
ResNet50	90.04	92.13	91.07
InceptionV3	91.31	93.02	92.16
VGG19	88.29	89.46	88.87

通过表2、图6、图7和图8可以看出,使用原始数据集训练模型的表现较差,使用了图像增广技术扩增数据集训练后,模型的整体性能有了一定的提高,精确率提高到84.67%;使用了迁移学习后,模型的性能大大提升,损失值迅速下降至0.0340, RMSE值下降至0.0574,精确率则提升到了96.24%;在此基础上扩增数据集使得模型性能再一次小幅度提升,损失值进一步下降至0.0326, RMSE值下降至0.0520,精确率提升至97.13%。而表3显示,经过重新测试后,与三组深度卷积神经网络模型进行对比得出,本文所构建的模型各方面的性能都比其他三组要好,识别的精确率更高。

最终,我们根据以上数据可以得出结论:比较模型的整体性能,扩增数据集+迁移学习>迁移学习>扩增数据集>原始数据集。利用图像增广技术扩充数据集能在一定程度上提升模型的性能,这是由于相比于原始数据集,模型得到了更加充分的训练。引入迁移学习技术能够大幅度提升模型的性能,这是由于迁移学习后使得模型在开始训练时就已经具备了一定的特征识别能力。值得一提的是,仅使用迁移学习的模型使用的是原始数据集且达到很好的效果,这说明迁移学习能够大幅减少训练所需样本数并且能够缩短训练时间。与其他网络模型对比也可看出,引入迁移学习的Xception网络更具有优势。

5 结论(Conclusion)

本文使用图像增广技术扩增数据集,利用迁移学习方法对Xception深层卷积神经网络模型进行训练,比较了模型前后的性能。根据实验分析发现:图像增广和迁移学习技术可以提升模型的性能,其中扩增数据集能够使得模型的训练更加充分,在一定程度上降低了模型的损失值和RMSE值,将模型

的精确率从79.82%提升到84.67%;而利用迁移学习技术使得模型的损失值和RMSE值大幅度下降,并且得到了96.24%的预测精确率,同时还缩减了训练模型的时间,与同级别的深层网络相比,在精确率、召回率和 F_1 指数等指标上都表现出了更好的性能。本文实验为深层卷积神经网络检测表面结构裂纹时出现的识别精度低、耗费时间长等问题提供了解决思路 and 解决方案。

参考文献(References)

- [1] KIM H, AHN E, SHIN M, et al. Crack and noncrack classification from concrete surface images using machine learning[J]. Structural Health Monitoring, 2019, 18(3):725-738.
- [2] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553):436.
- [3] ZHUANG F, QI Z, DUAN K, et al. A comprehensive survey on transfer learning[J]. Proceedings of the IEEE, 2020, 109(1):43-76.
- [4] JINSKUL N, TSAI C F, TSAI C E, et al. Enhancement of deep learning in image classification performance using xception with the swish activation function for colorectal polyp preliminary screening[J]. Mathematics, 2019, 7(12):1170.
- [5] XU H, SU X, WANG Y, et al. Automatic bridge crack detection using a convolutional neural network[J]. Applied Sciences, 2019, 9(14):2867.
- [6] 唐德可,王峰,王宏琦.基于标准化处理的手写符号快速识别方法[J].计算机应用与软件,2018,35(09):242-246.
- [7] 魏小娜,李英豪,王振宇,等.医学影像人工智能辅助诊断的样本扩增方法[J].计算机应用,2019,39(09):2558-2567.
- [8] OLIPHANT T E. Python for scientific computing[J]. Computing in Science & Engineering, 2007, 9(3):10-20.
- [9] WONGSUPHASAWAT K, SMILKOV D, WEXLER J, et al. Visualizing dataflow graphs of deep learning models in tensorflow[J]. IEEE Transactions on Visualization and Computer Graphics, 2017, 24(1):1-12.
- [10] 李清涛,任宇驰,王远,等.基于人工神经网络全连接层优化的线损异常诊断方法研究[J].电气应用,2020,39(04):82-88.
- [11] 周安众,罗可.一种卷积神经网络的稀疏性Dropout正则化方法[J].小型微型计算机系统,2018,39(08):1674-1679.
- [12] 谢璐阳,夏兆君,朱少华,等.基于卷积神经网络的图像识别过拟合问题分析与研究[J].软件工程,2019,22(10):27-29,26.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25:1097-1105.
- [14] 毕常遥,袁晓彤.基于Adam局部优化的分布式近似牛顿深度学习模型训练[J].计算机应用与软件,2021,38(10):278-283.

作者简介:

徐峥匀(1998-),男,硕士.研究领域:机器学习,深度学习.
钱松荣(1972-),女,博士,教授.研究领域:机器学习,深度学习.